



The 6th International Conference on Mining Science & Technology

Research on remote sensing data processing strategy and application based on grid operation

Zhu Hong-chun^{a,b}, Liu Hai-ying^c, Jiang Tao^a, Zhang Ji-xian^{c,*}

^aGeomatics College, SUST, Qingdao 266510, China;

^bThe Key Laboratory of Geomatics and Digital Technology, Shandong Province, Qingdao 266510, China;

^cChinese Academy of Surveying and Mapping, Beijing 100039, China;

^dCollege of Science, SUST, Qingdao 266510, China

Abstract

This research is satisfied with the needs of remote sensing massive data real time and fast transmission and processing. It developed the theory and methods for research of remote sensing processing based on grid operation. First, based on the generalized analysis to the concept and characters of RS grid and grid operation, the RS data grid processing characters were proposed. Second, taking the RS data grid processing basic strategy and the method as the guidance, a typical remote sensing data processing grid strategy and basic research methods were deal with, and specific tasks and data distribution strategy were put forward and used in the development of design grid model. Finally, the RS grid applications were completely deployed in the experimental grid environment and applied of which combined with experimental data. The operating results indicate that RS grid applications have higher efficiency and velocity.

Keywords: remote sensing; grid operation; processing; algorithm; application

1. Introduction

Recently, with sensors, platforms, data communication and related remote sensing technologies or theories continuously improved, the type and quantity of remote sensing data are expensing rapidly, particularly data quantity is grown geometrically. At the same time, in view of remote sensing data transmission and processing, more efficient ways have become the focus in RS research. Grid is the third tide of the internet and also is an integrated resource and service environment. Grid has the characters that can share resource and synergy with the work of various resources. So it can be achieved to maximize the sharing and application of the efficient realization of massive data and real-time processing. The grid operation is the joint operation pattern which is one kind of distributional way, the highly effective sharing data processing, the software and hardware use. The questions which

* Corresponding author. Tel: +86-13553065949; fax: +86-532-86057275.

E-mail address: zhc1977@gmail.com.

remote sensing data processing facing can be solved using the grid environment and the grid operator schema^{[1][2]}.

This article carried on the remote sensing data processing method research under grid environment. The operating procedure that RS data management integrated with the special processing was proposed. A typical remote sensing data processing grid strategy and basic research methods were dealt with, and specific tasks and data distribution strategy were put forward and used in the development of design grid model.

2. Grid operating and RS data processing

In the grid environment, the concrete data access, the data computation and the operator schema and the standard are different from traditional operation. The grid operating characteristics are resource sharing, joint operation, international open technical standard and dynamic.

RS processing is not only a special behavior, but also an analysis behavior that operator to carry on processing to the specific RS data using the method model. If the analyzing process is placed in the grid environment, it is called “the remote sensing processing grid”. Its function characteristic includes:

1. It is comprehensive resource pool integrated the RS data and processing resource;
2. The grid has realized separate grid resources connections through internet facility, for example computer, human, instrument, data, algorithm model and so on;
3. The grid has played the middleware character in view of logical constitution;
4. The resources integration management and what the grid realized were constituted the grid core part^{[3][4]}.

3. The basic strategies and methods of remote sensing data grid processing

3.1. Grid characteristic analysis of remote sensing data processing

Grid is a multi-machine collaborative computing environment, and parallel computing tasks are the significant feature of grid computing. According to the characteristics of image processing and the actual processing, we can sum up the digital image processing into the parallel geometry, adjacent parallel and parallel-bit pixel, the parallel operation of the four aspects of parallelism. In general, there are three models of remote sensing data-processing to achieve parallelism: ① stream line parallelism; ② functional Parallelism; ③ data parallelism. Due to the consistency and neighborhood characteristics of the image data, the idea of data parallelism is very natural, at the same time the parallel model is more suitable for the current mainstream of grid computing systems (such as Windows PC machine, Linux PC, etc.)^{[5][6]}.

3.2. The basic strategies and methods of remote sensing data processing based on parallelism

Remote sensing data grid parallel processing can be data in parallel and function in parallel. We can design the appropriate processing methods and strategies, according to block from the data and the division of tasks. The basic strategies and methods of remote sensing data grid processing are introduced here, combining the spatial domain image processing.

1. The image division by spatial domain

Adapting to the environment of multi-machines working together, the ideas of specific algorithm handling to block the spatial domain of remote sensing images are shown in Figure 1, the dashed lines in the figure indicated the remote sensing data processing course, which has not been devised.

In the strategy of blocking, remote sensing data processing tasks, which includes three stages, blocking, processing, and splicing turn into three stages by the task. The repeated numbers in the processing course depend on the numbers of image blocking to achieve a large amount of data of the computing tasks assigned to the grid in the balance. Applications of remote sensing data processing strategies are: spatial filtering, image of the geometric correction and radiant correction, un-supervised classification of the K-Mean classification^[7].

The calculation of total processing time using this method adapts to equation (1):

$$T_{all} = T_{cut} + \max(T_{proc}) + T_{merge} \quad (1)$$

In equation (1), T_{all} indicates total processing time, T_{cut} indicates image blocking time, $\max(T_{proc})$ indicates the max time of processing blocking image , T_{merge} indicates image mosaic time. Ignoring the time of the data blocking and image mosaic, we can assume that the performance of various computing nodes and configuration similar to the circumstances involved in n computer terms comparing with the stand-alone image processing, the amount of time is about $\frac{1}{n}$.

2. The division of processing steps
The specific tasks are divided into N-1 sub-steps and assigned N-1 computing nodes. After executing every sub-step, summary of the final steps is executed. The entire task completion time is divided into the max time of the sub-task and final step in processing time. Expressing as:

$$T_{all} = \max(T_{proc}) + T_{last-p} \tag{2}$$

In the equation (2), T_{all} indicates the total processing time, $\max(T_{proc})$ is the max time of N-1 steps, T_{last-p} indicates the final time of image synthesis. If the stand-alone machine execute the processing course, then the amount of time expressed as equation (3) (assuming the N nodes of the calculation is similar to the performance and configuration):

$$T_{all} = \sum_{i=1}^{n-1} T_{proc} + T_{last-p} \tag{3}$$

Compare equation (2) with (3), we can get the preliminary conclusion: the task of step-by-step implementation of multi-parallel processing consume much less time than that of stand-only processing, and the processing method approaches obviously higher efficiency.

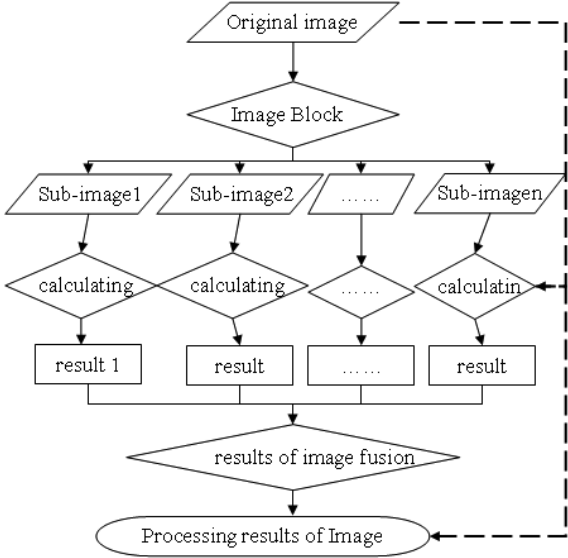


Fig. 1. The flow chart of RS data processing using data division

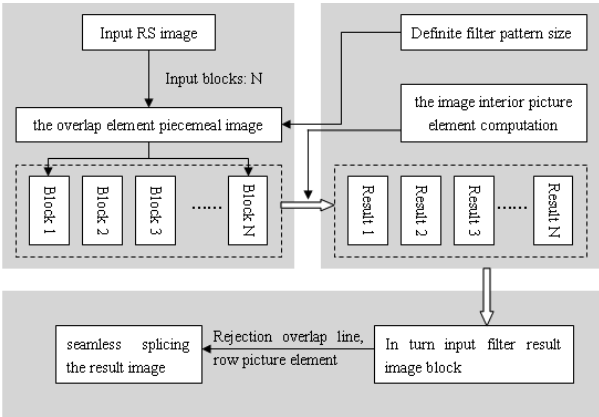


Fig. 2. The algorithm flow chart of filter processing based on image division method

4. Grid analysis algorithm design and application of specific remote sensing data processing

According to image division by spatial domain, the grid processing algorithm of image spatial filtering using sub-block strategy was designed and tested in grid environment. And according to image processing tasks division, the grid algorithm of regional automatic growth segmentation was designed and tested in grid environment. So, we can make out corresponding grid algorithm model.

4.1. The grid algorithm and flow of RS image spatial filtering processing

Spatial filtering processing is repeated iterative process of calculation pixel by pixel, and this course is simple, highly repetitive and the spatial correlation of calculation between pixels is also small (depending on the calculation of the size of the template). Therefore, in the light of data blocking, the grid used in the operation of spatial filtering algorithm described in the process is shown in Figure 2.

4.2. The grid algorithm and flow of RS image regional growing segmentation

1. The parallel analysis and strategy of image segmentation

(1) Algorithm parallelism: Regardless of what kind of growth segmentation criteria, the processes of seed points are targeted at the iteration process. So that parallel operation is possible.

(2) Segmentation parallel strategy: Various segmentation criteria can be integrated for use. Based on the growth of gray value segmentation results, the test for gray structural features and characteristics of the border region division can be simultaneously process to achieve better segmentation results.

(3) Seed growth parallelism: Selection of different seed carrying on the growth will be able to enhance the precision of division, but simultaneously will increase the computation burden. These burdens may be balanced and the high division precision can be achieved based on grid^[8].

So, a grid algorithm of regional growing segmentation was proposed.

2. Algorithm process

The concrete algorithm flow may describe as shown in Figure 3. This algorithm application was carried on in the grid environment. According to the computation seed picture, element integer assignment division treating processes number, synthesizes N parallel process and the division image, then completes the entire processing operation.

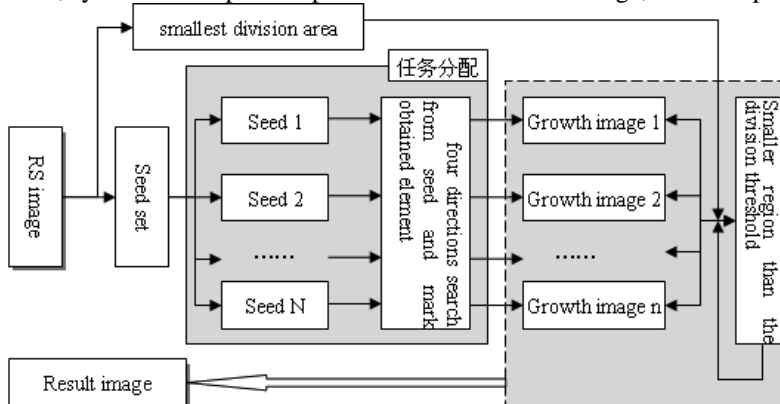


Fig. 3. The algorithm flow chart of image segmentation based on Task allocation

3. Some explanations

In order to avoid over-segmentation problems, we can make use of smoothing factor, for example Gaussian smoothing function, to preprocess the image in the segmentation process. We can utilize the mean gradient method to obtain the segmental threshold on the basis of the image preprocessing.

These steps of treating processes are a relapse replacing the growing point to carry on the growth. From the first step, the seed picture element sequences were established according to the respective gradation gap value Z. And finally, the preliminary image division was realized^[9].

4.3. The algorithm realizes and the grid application

In order to using the algorithm model at many machine or platform, the typical RS processing grid algorithm model were carried on and realized by unified procedure and execution environment. In this research, the pure C++ translating language was inherited from Linux machine. And other grid node machine deployed the GCC++ translation environment. This article partly proposed the typical RS processing algorithm sound code which mainly

uses the C++ language compilation, and EXE document which the production may carry out under the GCC++ translation environment.

The realized algorithm model contains the C code and the execution document. Meanwhile, they were resisted in the grid server as the grid resources by way of service issue description document statement in grid. The user may transfer these algorithm models when carries on concrete RS processing, and realizes typical remote sensing processing in grid. This article realizes two kind of typical RS processing algorithm model. Its name and the grid application method can be seen in Table 1.

Table 1. The tabulation of typical RS processing algorithm model and the application method

	Algorithm model	Grid application method
Spatial filtering processing	Image division procedure	The grid duty submission description script document contains three procedures, reads in a script document together, simultaneously contains to the task execution order stipulation parameter.
	Image filter procedure	
	Image merge procedure	
Image Segmentation by Region Growth	Growth seeds generate procedure	The grid duty submission description script document contains four procedures, reads in a script document together, simultaneously contains to the task execution order stipulation parameter. Duty which synchronizes carried on, must have the stipulation different growth picture element sources of information document parameter.
	Region growth procedure	
	Region increase procedure	
	Region merging procedure	

4.4. Grid implementation and performance evaluation for algorithm

In experiments, all grid computation nodes participated in processing completely and its execution process time was arranged in table 2 in order to use construction grid experiment environment.

Table 2. Remote sensing data grid process time

Processing duty	Processing mode	Excute time (second)
Spatial filter	Single processor	35
	Grid(6 nods)	8.48
Image segmentation	Single processor	89
	Grid (6 nods)	31

1. The single mode time is 35 seconds which is the longest for carrying on image filter processing. But the processing efficiency is also the highest in image piecemeal processing. Because when the piecemeal number is equal to the grid nodal point number, the grid resources used will achieve maximum limit.

2. The single mode time is 89 seconds which is the longest for carrying on image segmentations. The processing efficiency is the highest in grid environment when all computer nodes participate processing.

5. Conclusions

Satisfying the magnanimous RS resources joint operation and the data fast processing request using the grid technology, it is practical, feasible and quite complex topic. This article has conducted the thorough research on some questions about remote sensing processing grid operation. The concept and the characteristics of RS grid operation have been analyzed. The basic strategy and the method of RS data grid processing were proposed, and the RS data grid algorithm model was designed accordingly. Finally, the applications in grid experimental environment have been carried on.

The operating results for a comprehensive evaluation proved that the services example achieved higher efficiency

and better results. At the same time, the grid each kind of remote sensing resources combination used optimizes the key point and the man-machine interaction the grid operating mode discussion will be further studies.

Acknowledgements

This paper is supported by Open Research Found Program of the Key Laboratory of Geomatics and Digital Technology, Shandong Province (No.SD080711); “Taishan scholar” construction project special funds of Shandong Province; the National High Technology Research and Development Program (“863” Program) of China (No.200912Z12147).

References

- [1] Y.J. Zhang, Image Segmentation. Beijing: Science Press, 2001.
- [2] H. Zhu, Research on Griding Processing Strategies of Remote Sensing Image Segmentation by Region Growth. “Mapping without the Sun”——Techniques and Applications of Optical and SAR Imagery Fusion, 9 (2007) 288-290.
- [3] H. Zhu, Application of Distributed Processing to Remote Sensing Data Based on Condor. Journal of Shandong University of Science and Technology, 27 (6) 5-8.
- [4] Z.H. Yu, Y. Chen, P. Liu, Grid Computing. QingHua Press, Beijing, 2003.
- [5] J. Michael, Parallel Programming in C with MPI and OpenMP. QingHua Press, Beijing, 2004.
- [6] J. Smith, Close range photogrammetry for analyzing distressed trees. Photogrammetria, 42 (1987) 47-56.
- [7] N. Giovanni, Grid technology for the storage and processing of remote sensing data: Description of an application. Proceedings of SPIE - The International Society for Optical Engineering, 4881 (2002) 677-685.
- [8] J. Jeon, H. Kim, G. Y. Choi and H. Park, KAIST image computing system (KICS): A parallel architecture for real-time multimedia data processing. Journal of Systems Architecture, 46 (2000) 1403-1418.
- [9] S. Boussakta, A novel method for parallel image processing applications. Journal of Systems Architecture, 45 (1999) 825-839.